DOCUMENT RESUME

ED 317 548 TM 014 486

AUTHOR	Kerstiens, Gene
TITLE	A Slow Look at Speeded Reading Comprehension Tests.
PUB DATE	90
NOTE	21p.
PUB TYPE	Information Analyses (070) -- Journal Articles (080) -- Reports - Evaluative/Feasibility (142)
JOURNAL CIT	Review of Research in Developmental Education; v7 n3 1990
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Entrance Examinations; *Diagnostic Tests; Literature Reviews; Low Achievement; Postsecondary Education; Reading Comprehension; *Reading Tests; *Student Placement; Test Anxiety; *Test Use; Test Wiseness; *Timed Tests
IDENTIFIERS	*Speededness (Tests)

ABSTRACT

The definitions of and literature on test speededness are reviewed, and factors and areas in which the tests reveal their strengths and limitations are discussed. The focus of this article is on speeded reading comprehension tests used for entry or re-entry placement in postsecondary institutions. Speeded tests were prompted by the accelerating popularity of timed tests initially used for mass assessment of World War I recruits and conscripts. Early research reviews dealt almost entirely with studies based on comprehensive student populations. Later investigations began to direct attention toward underprepared populations. Almost inevitably, students who score low on the reading test portion of an entry assessment are candidates for intervention. Issues covered include speed set and anxiety, test wiseness and accuracy, and backwash. Speed set is a rate-of-work mind set purposefully executed to complete a task during the allotted time instead of the time the task demands. Backwash refers to the effect that a test of particular skill has on the acquisition of that skill. The preponderance of research indicates that the tests in question are inappropriate for comprehensive populations as well as developmental students. There are a number of acceptable alternatives to speeded comprehension tests. A 52-item list of references is included. (TJH)

A SLOW LOOK AT SPEEDED READING COMPREHENSION TESTS

By Gene Kerstiens

Review of Research in Developmental Education
Volume 7, Issue 3, 1990

National Center for Developmental Education
Appalachian State University
Boone, North Carolina 28608

A SLOW LOOK AT SPEEDED READING COMPREHENSION TESTS

By Gene Kerstiens

That placement testing is an experience facing nearly every entering freshman and many re-entry students is a well established postsecondary reality. The test or test battery typically includes a speeded reading comprehension component employing challenging reading passages that require exacting reading and thinking skills in order to answer an abundance of accompanying test items. So numerous are the passages and items and so limited is the testing time that many students do not complete the test unless they resort to testwise, score-inflating skimming or random responding strategies.

Almost inevitably, students who score low on the reading test portion of an assessment are candidates for intervention. Identified as developmental, remedial, underprepared, or otherwise, these students are counseled and/or required to participate in courses, programs or other treatments designed to remedy whatever reading deficiency or disorder the test can be interpreted to reveal. Having participated in a prescribed intervention and before gaining access to other programs or courses in the curriculum, these students once again

may be administered the same or another form of a speeded reading test. Unanimous research and expert opinion to the contrary, the sufficiently appreciated scores on this retest often are construed to predict competence in college-level reading skills. When we consider that at least half of the students evaluated in a given entering or re-entering college population can be identified as underprepared, the significance (or enormity) of the testing as well as the placement process takes on formidable proportions.

Addressing these matters, this investigation will explicate the definitions of test speededness, review the professional literature on the topic, and discuss identifiable factors and areas in which the tests reveal their strengths and limitations.

SPEEDED COMPREHENSION TESTS

Recognized in the literature is the distinction between power and speed tests. Generally, reading comprehenion tests having no time limits are characterized as power tests. When time is imposed so that all participants do not finish the test, it is considered a speeded instrument. Therefore, the degree of speededness assigned to a given test is relative to the completion rate or near completion rate of a norming population.

The Educational Testing Service "considers a test to be speeded if fewer than 100 percent of candidates reach 75 percent of the items and fewer than 80 percent of the candidates finish the test." (Davis, Kaiser, Boone 1987, p.9). Discussing

reasonable accommodation for developmental testing candidates, Morante (1989) asserts that 90 percent of students should complete all items; Nunnally (1978), identifying a "comfortable time limit" (p. 632), agrees. Studying developmental students' observed and stated behaviors as well as compared scores on nine standardized reading comprehension measures, Kerstiens (1986a and 1986b) used the term time-critical to identify any timed reading test scored without correction for error whose time constraints are such that it establishes or requires test-taking speed set by (1) encouraging the employment of score-inflating testwise strategies, (2) rewarding rapid responding at the expense of accuracy, or (3) eliciting a negative or interfering level of anxiety.

Among these and other numerous operational definitions that have evolved, there is little close concurrence (Davis, Kaiser, Boone 1987). Nonetheless, that the majority of reading comprehension tests presently implemented in assessment programs would qualify as highly speeded even under a conservative definitional umbrella seems reasonable.

REVIEW OF REVIEWS OF SEVENTY YEARS OF RESEARCH ON SPEEDED READING COMPREHENSION TESTS

Early-on and prompted by the accelerating popularity of timed tests initially used for mass assessment of World War I recruits and conscripts, Gates (1921) conducted an extensive review of research of reading tests. He learned that on timed

reading tests higher scorers were less accurate, passages represented a too limited number of academic content areas, and that chance-success responses inflated scores. Corroborating Gates' findings, Flanagan's (1939) bibliographical search as well as his own study also revealed that higher scoring students were less accurate than their lower scoring peers, that indeed students' scores were inversely related to their accuracy of response. Likewise, Carlson's (1949) review concluded that timed reading comprehension testing's psychometric shortcomings included the condition that they penalized slower but more accurate readers. In 1958, Letson reviewed nineteen studies stating that "speed of working" unfairly escalated scores, particularly on instruments presenting more difficult test items. Rankin's (1962) investigation cited numerous critics who found design flaws in time-critical tests of reading comprehension, and Davis (1962) discovered that the same tests, scored without correction for chance success, were invalid. Both of the writers agreed that post-test gains on timed reading tests were enjoyed chiefly by sacrificing accuracy of comprehension. Taking previous studies and reviews into account together with the findings of his own study, Tillman (1977), in a seemingly fin de siecle pronouncement, declared that "Ironically, the increasing popularity of certain tests seems to be inversely related to the negative comments of critics." (p. 253)

While earlier research reviews dealt almost entirely with studies based upon comprehensive student populations, later investigations began to direct their attention to underprepared

populations. This area of investigation apparently had gone unrecognized, having provoked insufficient academic curiosity to inspire publishable research. Focusing on the experience of less academically adept students and citing ten relatively recent studies centered on the use of the most frequently employed test, Perkins (1984) strongly questioned its validity. She recommended "a re-examination of the use of the NDRT as an instrument which determines whether or not college freshmen possess adequate reading skills." (p. 67) Rounds' (1984) review echoed this dissatisfaction with the NDRT and other popular reading comprehension tests, finding them to be poor assessment instruments as well as unreliable predictors of California community college course success. Kerstiens' (1986ab) studies and research reviews concluded that developmental students' scores on time-critical reading test measures were attributable largely to student rate of response, accuracy of response remaining at about thirty-five percent irrespective of score level. Finally, a carefully focused investigation involving an underprepared Tennessee population together with an extensive research review (Davis, Kaiser, Boone 1987) recommended that time limits for the test in question be extended by thirty percent to assure fairer measurement.

Truly, among the seven decades of studies and statements on the topic of comprehension test speededness, there are a handful of publications affirming that, given the properties of certain instruments and the level of the testing population, the negative effects of speededness are minimal, inconsequential or inoffensive. None of these champions the condition of timing,

whose apparent virtues are administrative convenience and the sorting of a talented reading elite. (Carver, 1984; Cummins 1981; Raygor, 1978.) As shall be demonstrated, the psychometric, behavioral, affective, and curricular side effects of reading comprehension test speededness are substantial.

SPEED SET AND ANXIETY

Speed set is a rate-of-work mind set purposefully executed to complete a task during the allotted time instead of the time the task demands. Although virtually all testing consists of less than life-like samplings, time-critical tests effect a sprinted setting that is considered demonstrably artificial. (Bartin 1975; Rubenstein, Kender, Mace 1988) Nor do they address themselves precisely to assessing the basic abilities which facilitate a student's functioning while learning from challenging textual material. (McDonald 1966; Tuinman 1970) Especially for students given to deliberative reading, the urgency of the testing situation is not consistent with work requiring a study-reading approach rather than adroitly executed skim-and-respond strategies. (Rodriguez 1985; Townsend 1965) Dealing with difficult material, even the best readers retard rates in order to re-investigate or reflect. (Carver 1985) But the typical underprepared student often thinks through silent speech (Pauk 1968; Feldman 1976) and routinely exercises re-reading behaviors while striving to comprehend. (Van Blerkom and Van Blerkom 1981) Also, the test items themselves often require more analysis, manipulation and interpretation than the reading

passages they attend (Raygor 1978; Raygor and Flippo 1971) -- all of these tasks requiring careful reading and some reflection if the intention is to comprehend accurately.

Closely related to the generic speed set issue is the phenomenon of anxiety. Virtually all writers agree that the imposition of time limits heightens anxiety in subjects and that the higher the speededness factor, the greater the anxiety. Wark and Bennett (1981) and Wark, Testenrud and Nelms (1972) studied the psychological as well as physiological aspects of anxiousness under speeded reading test conditions to establish that higher-anxious students' scores are negatively affected. Cummins (1981) and Davis, Kaiser and Boone (1987) indicate that such over-anxious reactions inhibit concentration. In their Miami-Dade study (1988), Schinoff and Steed surveyed and observed developmental students to learn that they thought the speededness factor was undesirable. As early as 1951, Preston and Botel declared that the speed set/anxiety factor rendered timed reading tests "invalid and untenable." (p.71)

TESTWISENESS AND ACCURACY

"Testwiseness is the ability to correctly answer test questions on some basis other than knowledge that the questions were designed to measure." (Ferrell 1977, p. 138) Since the majority of timed reading comprehension tests employed in postsecondary assessment are scored without correction for error (Davis 1962), they are particularly vulnerable to score-inflating test-taking strategies. (Gates 1921; Kerstiens 1986ab; Morante 1989; Townsend 1965) The more challenging the test's reading

passages and test items, the more constrained the time limit, and the slower the rate of response of the student, the more likely the testwise candidate will resort to cultivated guessing and random responding strategies. To the extent that students have knowledge of these strategies and are willing to exercise it, their test scores can be significantly enhanced, these practices diminishing the integrity of individual scores and the validity of the instrument. Of course, non-testwise students or those whose ethical predisposition prevents them from exerting their testwiseness are penalized.

That test speededness affects a student's rate and accuracy of response is well supported in the literature. (Carlson 1949; Cronan 1987; Gabriel and Richards 1988; Jolly 1985; Sadler and Whimbey 1985; Stetson 1981; Townsend 1965) And that students can learn or be taught guessing strategies to increase their scores is commonly acknowledged. But that students actually improve their reading comprehension having applied these strategies is questionable and undocumented.

Therefore, writers express concern that encouraging test-taking guessing habits fosters certain inappropriate learning attitudes and behaviors: wild guessing (Gates 1921); answer grabbing (Sadler and Whimbey 1985); one-shot thinking (Gibson 1986); chance success attitude (Davis 1962); and skim and guess syndrome. (Townsend 1965). Accordingly, strongly suggested in the literature is the notion that utilizing instruments that encourage these behaviors, and, further, drive the direction of the reading curriculum designed to inculcate effective reading

skills (1) encourage inappropriate habits and attitudes toward learning, (2) discourage the sharpening of critical reading and thinking skills, and (3) essentially deter students from the realities of effectively learning from their textbooks.

BACKWASH

"Backwash refers to the effect a test of particular skill has on the acquisition of that skill." (Backwash 1983, p. 1.) The backwash result can produce desirable or undesirable effects; undesirable conditions occur when there is a mismatch between measurement and instruction. (Popham 1978) Researchers concur that the fall-out of time-critical reading comprehension testing is short of salutary.

Besides having some unrewarding immediate influences on students' reading behaviors, our typical reading assessment practices may also provide students with unrealistic expectations and may influence practitioners to misdirect instruction. Since these tests do not present a fair sampling of a broad range of subject matter that represents a variety of academic disciplines (Morante 1989; Wood 1989), they cannot guage the scope and scale of students' schema or prior knowledge that affect comprehension. (Durkin 1981; Hirsch 1987) Nor are the tests able to measure skills needed to cope with the wide range of difficulty encountered in assigned textbooks. (Bartin; 1975 Kurak 1967) Because many of these tests are freighted with "bookish" (Raygor 1978, P. 1210) passages and test items that appear to be calculatedly intimidating, they fail to "discriminate best among students of

low proficiency" (Morante 1989, P.2) whose need for accurate assessment is greatest. Thus students as well as practitioners are presented with tests suggesting directions for skill attainment that are unrealistic and ultimately frustrating.

But of particular concern of late is this testing's backwash effect on the reading curriculum. Apprehension emerges from the notion that if testing is allowed to drive the reading curriculum "students will not learn the skills and abilities that they need to perform college reading tasks effectively." (Wood 1988, p. 229) Wood (1989) is further disquieted by the expectation that if tests do not authentically reflect the demands of college reading assignments, they tend to impose "artificial limits" on the reading curriculum.

The concern is hardly trivial. For in her celebrated and repeatedly cited research, Fairbanks (1972) (also see Farrell 1975) specifies the reading skills whose mastery predicts success in college courses. We know that the academic achievement of college students improves when they learn to identify main ideas, distinguish between fact and opinion, and recognize and interpret inferences while managing a range of metacognitive sub-skills enumerated by Mickler (1989). Any notion that the tests in question begin to evaluate these skills fairly is unreasonable. That college reading programs can prepare their underprepared students while focusing on the skills measured by the tests is unjustifiable.

DISCUSSION

Of course, there are assessment alternatives to speeded

reading comprehension tests. (Allina 1987· Schinoff and Steed 1988) And given the preponderence of research indicating that the tests in question are inappropriate for comprehensive populations as well as developmental students, one wonders how the academic community can continue to ignore the findings with seemingly serene indifference. Perhaps the poignant question is How can those responsible for postsecondary assessment continue to sacrifice accuracy of measurement for administrative expedience?

But that is the topic of another publication.

REFERENCES

Allina, A. (1987). Beyond standardized tests: Admissions alternatives that work. Cambridge, MA: National Center for Fair and Open Testing. (ERIC Document Reproduction Service Report No. 292 812).

Backwash (1983). ILC Testing Newsletter, 1, 1-2.

Bartin, N. (1975) Standardized reading testing in college: Some cautions. In R. Sugimoto, Ed., Eighth Proceedings of the Western College Reading Association (pp. 31-34). Whittier, CA: The College Reading and Learning Association.

Carlson, T. (1949). The relationship between speed and accuracy of comprehension. Journal of Educational Research, 42, 500-512.

Carver, R. (1984). Rauding theory predictions of amount comprehended under different purposes and speeded reading conditions. Reading Research Quarterly, 29, 205-218.

Carver, R. (1985). How good are some of the world's best readers? Reading Research Quarter /, 20, 389-419.

Cronan, T. (1987). Reading rate: The measurement dilemma revisited. Forum for Reading, 19, 31-37.

Cummins, R. (1981). Test review: The Nelson-Denny Reading Test (Forms E and F). Journal of Reading, 25, 54-59.

Davis, F. (1962). Measurement of improvement in reading skills

courses. In E. Bliesmer and R. Staiger, Eds., Eleventh Yearbook of the National Reading Conference (pp. 30-40). Boone, NC: The National Reading Conference.

Davis, T., Kaiser, R, and Boone, J. (1987). Speededness of the Academic Assessment Placement Program (AAPP) reading comprehension test. Nashville, TN: Board of Regents of the State University and Community College System of Tennessee.

Durkin, D. (1981). What is the value of the new interest in reading comprehension? Language Arts, 58, 23-43.

Fairbanks, M. (1977). An analytical study of the relationship of specified features of reported college reading improvement programs to program effect on academic achievement. Dissertation Abstracts International, 33 6114A. (University Microfilms, Order No. 73-12,938).

Farrell, T. (1975). Reading in the community college. College English, 37, 40-46.

Feldman, J. (1976). Why I move my lips when I read. In M. Douglass, Ed., Fortieth Yearbook of the Claremont Reading Conference (pp. 128-134). Claremont, CA: The Claremont Graduate School.

Ferrell, G. (1977). Development of a test of test-wiseness. In G. Enright, Ed., Tenth Proceedings of the Western College Reading Association (pp. 138-142). Whittier, CA: The College Reading and Learning Association.

Flanagan, J. (1939). A study of the effect of comprehension of varying speeds of reading. In Research in the Foundations of American Education (pp. 47-50). Washington, D.C.: American Educational Research Association.

Gabriel, D. and Richards, I. (1988). Vocabulary, intelligence and reading comprehension. Parma Heights, OH: Cuyahoga Community College. (ERIC Document Reproduction Service Report No. 301 248).

Cates, A. (1921). An experimental and statistical study of reading and reading tests. Journal of Educational Psychology, 12, September, October, November, 303-314, 378-391, 445-464.

Gibson, S. (1986). A three-dimensional approach to test-taking instruction. Journal of College Reading and Learning, 19, 51-58.

Hirsch, E. (1987). Cultural literacy: What every American needs to know. Boston: Houghton-Mifflin.

Johnston. P. (1984). Prior knowledge and reading comprehension test bias. Reading Research Quarterly, 19, 219-239.

Jolly, S. (1985). The effect of test speededness and random guessing on the validity of reading comprehension scores. Paper presented at the meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service Report No. ED 262 063).

Kerstiens, G. (1971). Directions for research and innovation in

junior college reading programs. Topical Paper No. 18, ERIC Clearinghouse for Junior Colleges, Los Angeles: UCLA Graduate School of Education and the University Library.

Kerstiens, G. (1986a). A testimonial on timed testing: Developmental students and reading comprehension tests. In M. Douglass, Ed., Fiftieth Yearbook of the Claremont Reading Conference, (pp. 261-268). Claremont, CA: The Claremont Graduate School.

Kerstiens, G. (1986b). Time-critical reading comprehension tests and developmental students. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service Report No. 278 700).

Kurak, A. (1967). Developing a junior college reading test. The General Education Sounding Board, 4, 1-7.

Letson, C. (1958). Speed of comprehension in reading. Journal of Educational Psychology, 52, 49-53.

McDonald, A. (1971). Measuring reading performance. In G. Schick and M. May, Eds., Fifteenth Yearbook of the National Reading Conference, (216-221). Milwaukee, Wisconsin: The National Reading Conference.

Mickler, J. (1989) Comprehending textbooks: Building a relationship between texts and readers. Review of Research in Developmental Education, 7, (1), 1-4.

Morante, E. (1989). Selecting tests and placing students. Journal of Developmental Education, 13, 2-6.

Nunnally, J. (1978). Psychometric means. New York: McGraw-Hill.

Pauk, W. (1968). Skills beyond the reading skills. In F. Greene, Ed., Proceedings of the Tenth Annual University of Pittsburgh Reading Conference, No. 6 (pp. 127-140). Pittsburgh, PA: The University of Pittsburgh Reading Conference.

Perkins, D. (1984). Assessment of the use of the Nelson-Denny Reading Test. Forum for Reading, 15, 64-69.

Popham, J. (1978). Modern measurement technology: Miracle or mirage? In G. Enright, Ed., Eleventh Proceedings of the Western College Reading Association (pp. 12-14). Whittier, CA: The College Reading and Learning Association.

Preston, R. and Botel, M. (1951). Reading comprehension under timed and untimed conditions. School and Society, 74, 71.

Rankin, E. (1962). The relationship between reading rate and comprehension. In E. Bliesmer and R. Staiger, Eds., Eleventh Yearbook of the National Reading Conference (pp. 1-5). Boone, NC: The National Reading Conference.

Rodriguez, J. (1985). When reading less can mean understanding more. Journal of Reading, 28, 701-705.

Rounds, J. (1984). Assessment and placement in language arts courses: Three major tests. Journal of College Reading and Learning, 17, 94-100.

Raygor, A. (1978). The Nelson-Denny Reading Test (Form C and D). In O.K. Buros, Ed., Eighth Mental Measurements Yearbook, Vol. 2, (1209-1211). Highland Park, NJ: The Gryphon Press.

Raygor, A. and Flippo, R. (1981). Varieties of comprehension measures: A comparison of intercorrelations among several reading tests. In G. McNinch, Ed., Comprehension: Process and Product. Hattiesburg, Mississippi: American Reading Forum. (ERIC Document Reproduction Service Report No. 198 485).

Rubenstein, H., Kender, J., and Mace, C. (1988). Do tests penalize readers for poor short term memory? Journal of Reading, 32, 4-10.

Sadler, W. and Whimbey, A. (1985). A holistic approach to improving thinking skills. Phi Delta Kappan, 67, 199-203.

Schinoff, R. and Steed, L. (1988). The Computerized Adaptive Testing Program at Miami-Dade Community College, South Campus. In Computerized Adaptive Testing: The State of the Art in Assessment at Three Community Colleges (pp. 25-36). Laguna Hills, CA: The League for Innovation in Community Colleges.

Stetson, E. (1982). Reading tests don't cheat, do they? Journal of Reading, 25, 634-639.

Tillman, C. (1977). Readability and other factors in college

reading tests: A critique of the Diagnostic Reading Test, the Nelson-Denny Reading Test, and the McGraw-Hill Basic Reading Test. In D. pearson and J. Hanks, Eds., Twenty-sixth Yearbook of the National Reading Conference (pp. 253-258). Rochester, NY: The National Reading Conference.

Tuinman, J. (1970). Aspects of the assessment of acquisition of information from reading passages. Unpublished Doctoral Dissertation, University of Georgia.

Townsend, A. (1965). The Nelson-Denny Reading Test: Vocabulary-comprehension-rate. In O.K. Buros, Ed., The Sixth Mental Measurements Yearbook (pp. 1078-1080). Highland Park: NJ: The Gryphon Press.

Van Blerkom, M. and Van Blerkom, D. (1981). Misleading findings from the Nelson-Denny Reading Test. Journal of Reading, 24, 736-737.

Wark, D. and Bennett, J. (1981). The measurement of test anxiety in a reading class. Reading World, 20, 215-222.

Wark. D. Testenrud. D. and Nelms, G. (1972) Heart rate and reading. In F. Greene, Ed., Twenty-first Yearbook of the National Reading Conference, (pp. 153-160), Milwaukee, Wisconsin: The National Reading Conference.

Wood, N. (1988). Standardized reading tests and the postsecondary reading curriculum. Journal of Reading, 3, 224-230.

Wood, N. (1989). Reading tests and reading assessment. Journal of Developmental Education, 13, 14-16, 18-19.

ACKNOWLEDGEMENT: Gene Kerstiens is Director of The Learning Assistance Center, Scottsdale Community College, and Director of Andragogy Associates, College-Adult Learning Specialists.
